The use of canonical analysis

Kuylen, A.A.A.; Verhallen, T.M.M.

Published in:
Journal of Economic Psychology

Publication date:
1981

Journal of Economic Psychology 1 (1981) 217–237
North-Holland Publishing Company

THE USE OF CANONICAL ANALYSIS

Anton A.A. KUYLEN and Theo M.M. VERHALLEN *

Tilburg University, The Netherlands

Received August 25, 1980; accepted April 20, 1981

The first part of this non-technical review of canonical analysis is concerned with the principle, the data requirements, the interpretation, the evaluation and the application in consumer research of canonical analysis. In this context, concepts as canonical correlation, weights, loadings and scores are explained. Their interrelationships are discussed. Then the characteristics of specific forms of canonical analysis: canonical correlation, canonical regression, redundancy analysis and partial canonical analysis are discussed. Their different application properties are emphasized. The last part consists of an application of canonical correlation analysis for brand positioning. Here the relationship with discriminant analysis is illustrated. Furthermore the usage of canonical analysis for optimal scaling purposes is illustrated for the same example.

1. Introduction

This article aims at making the technique of canonical analysis accessible to the researcher in the consumer behavior area. In the first part a general review is given of the essentials, the data requirements and the interpretation and evaluation of problems associated with canonical analysis. This is followed by a short discussion of specific forms of canonical analysis and applications in consumer research. In the last part, an example on brand positioning is given to demonstrate both the relationship of canonical analysis with discriminant analysis as well as its use for optimal scaling purposes.

* The authors wish to thank Meeuwis J. Geleijns for programming a package on canonical analysis and his useful comments on this paper.
Mailing address: A.A.A. Kuylen, Katholieke Hogeschool, G 712, Postbus 90153, 5000 LE Tilburg, The Netherlands.

1.1. *What is canonical analysis?*

The main characteristic of canonical analysis is the investigation of the relationship between two sets of variables. One set is the predictor set or, analytically, the set of independent variables. The second consists of the criteria or the dependent variables. A frequently occurring research problem with this structure is when the relationship between attitudes, the predictors, and product-usage characteristics, the criteria, is to be investigated. In a canonical analysis, variates are computed from both sets of variables. A canonical variate is similar to a factor in a principal component analysis, with the difference that a variate consists of a maximally correlated predictor and a criterion part. Analogous to factor analysis, a maximum of N variates (factors) can be extracted, which are independent of each other. N is the number of variables from the smallest set.

To get acquainted with canonical analysis an example as we would obtain from many computer programs (*e.g.* SPSS) is given in table 1.

Table 1
An example of canonical analysis.

Variables	Variate 1 [a] can. weights	Variate 2 [a] can. weights
Predictor set		
Extensive	0.21	0.17
Sober	−0.04	0.13
Imaginative	0.10	0.01
Varied	0.30	−0.17
With family	−0.07	0.09
With care	−0.09	0.22
Fast	−0.29	−0.45
Neat	−0.02	0.41
Good looking	−0.18	−0.19
Criterion set		
Bread	0.23	−0.32
Meats	0.65	−0.33
Cheese	0.08	0.05
Eggs	0.15	0.03
Table cloth	0.15	0.69
Dishes	−0.10	0.33
Canonical correlation	0.95	0.80

[a] The canonical variates are standardized.

In this example the relationship between breakfast evaluations and the usage of breakfast products has been investigated using canonical analysis. The first two canonical variates are given here.

Canonical correlations (0.95 and 0.80) are analogous to ordinary correlation coefficients. It is important to keep in mind that a canonical correlation expresses the association between two underlying constructs. The relationships between the original observed variables cannot be derived directly from this. The implication of this will be discussed in the following section.

1.2. The interpretation of canonical analysis results

For the user of canonical analysis a number of problems may arise with regard to the interpretation and evaluation of results obtained from canonical analysis.

1.2.1. Canonical weights

The canonical weights are comparable with *b*-weights or, when the variates have been standardized, as usually is done in computer programs,with beta-weights from a multiple regression analysis. They serve to transform the original variables in such a way as to obtain a maximum correlation between predictor and criterion sets of variates. The magnitude of a weight expresses the importance of a variable from one set with regard to the other set in obtaining a maximum correlation between sets. We could start from these weights to interpret the results. However there are two problems:

(1) In the first place these weights may be unstable due to multicollinearity [1]. Some variables may obtain a small weight or even a negative weight because of the fact that the variance in a variable has already been explained by other variables. In this type of situation the weights do not give a clear picture of the relevance of the variables.

(2) Canonical analysis maximizes the correlation between predictor and criterion parts of significant canonical variates. The correlation, however, is computed on constructed, not observed variables. Technically, a canonical solution is the maximum correlation between pairs of

[1] Multicollinearity exists when variables from the same set are mutually highly correlated. When one wants to get a picture of the size of the multicollinearity, the coefficient of multiple determination (R^2) can be computed for each variable with all others in the same set. See Johnston (1972).

linearly transformed variables. As many new orthogonal variates may be computed from the residual variance as the number of variables from the smallest set. In table 1, with six criterion variables a maximum of six orthogonal canonical variates may be computed. Every canonical variate is the regression of the constructed, non-observed variate on the observed variables. As these are constructed variables it is not necessary that the relationship between actual and observed variables has any significance.

In order to investigate how much of the variance of the observed variables is retained in a canonical variate, it is necessary to compute canonical loadings. A canonical loading, analogous to factor loadings in a factor analysis, expresses the association between a variable and a canonical variate. Thus it offers an opportunity for the interpretation of a variate.

1.2.2. Canonical loadings

The canonical loadings can be found by correlating the raw variable scores with the variate scores. These canonical variate scores are analogous to factor scores in factor analysis. They express the scores of respondents (or objects on which the variables are measured) on the canonical variates. An example of the computational procedure involved is given below. Thus the correlation of variable X_i with F_i forms the canonical loading of variable X_i on variate F_i. In table 2 the canonical variate loadings for our example are presented.

One interpretation of the results based on the canonical loadings is as follows: persons who evaluate their breakfast as varied, extensive and imaginative tend to eat meats, bread and cheese at breakfast (variate 1), and people who consider their breakfast as neat, good looking, prepared with care tend to use table a cloth and dishes when having breakfast.

Although the structure of the loadings in this example demonstrates some similarity with those of the weights, there are important differences due to multicollinearity. This means that a researcher when taking the canonical weights for interpreting the content of the canonical variates, would be in error. The percentages of explained variance, 18.7 and 14.7% for the criterion variables, and 15.7 and 19.6% for the predictor variables, express the relationship between actual observed variables with the underlying, constructed canonical variates.

Table 2
A completed example of canonical analysis.

Variables	Variate 1		Variate 2	
	Canonical loadings	Canonical weights	Canonical loadings	Canonical weights
Predictor set				
Extensive	0.53	0.21	0.35	0.17
Sober	−0.43	−0.04	0.00	0.13
Imaginative	0.50	0.10	0.24	0.01
Varied	0.69	0.30	0.15	−0.17
With family	0.14	−0.07	0.53	0.09
With care	0.24	−0.09	0.58	0.22
Fast	−0.26	−0.29	−0.43	−0.45
Neat	0.08	−0.02	0.66	0.41
Good looking	0.26	−0.18	0.57	−0.19
Explained variance	15.7%		19.6%	
Criterion set				
Bread	0.48	0.23	0.12	−0.32
Meats	0.72	0.65	0.06	−0.33
Cheese	0.45	0.08	0.21	0.05
Eggs	0.39	0.15	0.30	0.03
Table cloth	0.13	0.15	0.72	0.69
Dishes	0.05	−0.10	0.46	0.33
Explained variance	18.7%		14.7%	
Canonical correlation	0.95		0.80	
Redundancy (y/x)	0.17		0.09	

1.2.3. Rotation

As with factor analysis, the matrix of loadings (or weights) can be rotated. Rotation may be important for two reasons. In most cases, as for instance with varimax rotation, the rotation leads up to a simpler structure with better interpretable results. A second consequence of rotation is factorial invariance. This means that the results of a solution can be generalized. A solution is said to be invariant when the same groups of variables are found repeatedly on variates whenever at least some important variables, which reflect the underlying construct, are used in the repeated analysis. An unrotated variate solution is more than a rotated solution dependent on all variables involved in the analysis. As with factor analysis it is up to the researcher to decide which solution is preferred. The predictor and criterion parts of variates can be rotated simultaneously as well as separately. In the latter case, however, it is possible that the initial solution is changed completely. With simultaneous rotation the canonical correlations will be more evenly spread over the different variates. The canonical loadings in our example are found after varimax rotation.

1.2.4. Canonical variate scores

Canonical variate scores express the scores of respondents on the canonical variates. They are found by multiplying the z-scores of respondents with the canonical weights. The computation for respondent one on the first canonical variate is given in table 3. The variate score for the first respondent on the criterion part of the first variate is 1.09, the variate score on the corresponding predictor part is 0.73. These variate scores can be used for subsequent analysis. If we compute for all respondents their scores on predictor and criterion parts of the variates, the correlations between them are equal to the canonical correlations.

1.3. The evaluation of a canonical analysis solution

Whenever a high canonical correlation between pairs of variates is found, this does not necessarily mean that the canonical analysis yields a useful and interpretable solution. When only one or a few variables are highly correlated with the canonical variate, and thus show high loadings, the total amount of variance explained in the observed variables will be low. Then the canonical structure indicates only a specific relationship between a few predictor and criterion variables. Then no relationship between more general constructs exists. Thus, a first additional measure to evaluate a canonical analysis solution is the *proportion of explained variance.* This may be computed from (1).

$$R_Y^2 = \frac{\sum_{j=1}^{n} (CL_j)^2}{n} \tag{1}$$

in which R_Y^2 = the proportion variance explained in the y (criterion) set;
CL_j = the canonical loadings in the y-set;
n = the number of variables in the y-set.

In a second measure, the *redundancy coefficient*, the two characteristics of a canonical solution, the canonical correlation and the amount of explained variance, are both taken into account. This measure expresses the amount of explained variance in one set, given the other set. The redundancy can be computed from (2).

$$R^2(y/x) = R^2_{c.k} \frac{\sum_{j=1}^{n} (CL_j)^2}{n} \tag{2}$$

in which $R^2\ (y/x)$ = the redundancy in y (the criterion set) given x (the predictor set);
$R^2_{c.k}$ = the squared canonical correlation of the K-th pair of variates;
n = the number of variables in the criterion set.

When we substitute the values from table 2 into formula (2) we find the redundancy in the criterion set given the predictor set. For the first canonical variate:

$$R^2(x/y) = 0.95^2\{(0.53)^2 + (-0.43)^2 + (0.50)^2 + \ldots$$

$$+ (0.08)^2 + (0.26)^2\} = 0.14$$

and

$$R^2(y/x) = 0.95^2\{(0.48)^2 + (0.72)^2 + (0.45)^2 + (0.39)^2$$

$$+ (0.13)^2 + (0.05)^2\} = 0.17$$

Then it is possible to compute the *proportion of redundancy* in one set given the other according to (3).

$$V_k(y/x) = \frac{R_i^2(y/x)}{\sum_{i=1}^{k} R^2(y/x)} \tag{3}$$

in which $V_k\ (y/x)$ = the proportion of redundancy in the y set given the x set;

$R_i^2\ (y/x)$ = the percentage explained variance in the y set, given the x set (formula 2);

k = the number of variates;

$\sum_{i=1}^{1} R^2\ (y/x)$ = the sum of percentages explained variance in the y set given the x set, with k canonical variates.

This measure expresses the amount of the total redundancy accounted for by each y variate. When there are more significant canonical variates the decrease in the proportions of redundancy tells us how many canonical variates to accept.

In summary, the canonical loadings are important for the interpretation of the canonical variates. With help of the canonical loadings we can get an impression of the importance of a canonical variate by means of the percentage explained variance, the redundancy and the proportion of redundancy. For a proper interpretation of canonical solution, we recommended that one inspects both the canonical weights

Table 3
Computation of variate scores on the first predictor and criterion variates for one respondent.

Variables	Canonical weights	×	Z-scores	=	Variate score
Predictors					
Extensive	0.21		1.20		0.25
Sober	−0.04		−0.35		0.01
Imaginative	0.10		0.75		0.08
Varied	0.30		−0.10		−0.03
With family	−0.07		1.12		−0.08
With care	−0.09		−1.60		0.14
Fast	−0.29		−0.75		0.22
Neat	−0.02		0.19		−0.00
Good looking	−0.18		−0.80		0.14
					$(\Sigma=)X_1=0.73$
Criteria					
Bread	0.23		1.50		0.35
Meats	0.65		0.95		0.62
Cheese	0.08		−0.35		−0.03
Eggs	0.15		0.63		0.10
Table cloth	0.15		−0.37		−0.06
Dishes	−0.10		−1.12		0.11
					$(\Sigma=)Y_1=1.09$

and the corresponding canonical loadings. Canonical weights provide an insight into the predictive qualities of the variables, while the canonical loadings are necessary for the interpretation of the nature of the relationship. Large differences between weights and loadings (absolute and in direction) can provide indications for moderator and suppressor variables as well as for non-linear relationships. See Schaninger *et al.* (1980) for a more detailed discussion

1.4. When to use canonical analysis?

Canonical analysis is a recommended technique for analyzing several predictor variables and criterion variables simultaneously. Especially when the criterion variables are mutually correlated, a canonical analysis is appropriate. In such a case complex relationships between structures in predictor and criterion variables may be found.
From the above example it appears that canonical analysis is both a structural and a functional technique: the predictor and criterion set are structured in such a way as to create a maximum correlation between sets.
Separate multiple regression analyses for each of the criterion variables would neglect the interrelations of the criteria, while factor analyses on each of the two sets of variables would neglect the relationships between predictors and criteria. Correlations between predictor and criterion factors obtained from factor analyses would never be as high as between variates found from canonical analysis.
An example of the superiority of canonical analysis over factor analysis is given by Wendt (1979).

1.5. Data requirements for canonical analysis

For descriptive use of canonical analysis it suffices that predictor and criterion variables are dichotomous or of interval level. When one also wants to test the significance of the relationships between variates, the requirements of multivariate normality and homogeneity of variance should be met.

1.6. Applications of canonical analysis

The first applications of canonical analysis in consumer research were concerned with classical themes. Sparks and Tucker (1971) investigated the relationship between personality traits and product usage and Alpert (1971) investigated the relationship between personality traits and automobile choice. Baumgarten and Ring (1971) and Darden and Reynolds (1971) investigated the relationship between socio-demografics and media usage, respectively, shopping behavior and product buying the new technique is used to re-evaluate classical problems. For practice oriented market researchers the applications of Frank and Strain (1972) and Fornell and Westbrook (1978) are interesting. They use canonical analysis to perform a market segmentation. Using panel data Frank and Strain (1972) cluster respondents on the variate scores obtained from the relationship between personal characteristics and product usage variables. Frank and Strain (1972) use the variate scores on the predictors, the personal characteristics, for segmentation while Fornell and Westbrook (1978) perform a segmentation based on the criterion variate scores found in a canonical analysis on personal and decision-process characteristics with information usage variables. Both market segmentation applications illustrate that the requirements regarding the predictive and discriminative power of market segments form an integral part of a canonical analysis approach.

A different kind of application is from Carmone (1977) who used the weights found in canonical correlation to determine (cross) price elasticities. The applications described here illustrate the point stated before that canonical analysis can be alternative for both functional (*e.g.* regression-) as well as structural (*e.g.* factor-) analyses. This point is elaborated further in the example on brand positioning presented below.

2. Types of canonical analysis

Thus far a distinction has been made between two sets of variables: the predictor and the criterion set. These sets might have been reversed without any consequences for the computational procedure and interpretation of the results. So, implicitly, the discussion was focused on the symmetrical form of canonical analysis, canonical correlation, that may be considered as the original and major form of canonical analysis. In the following, the asymmetric forms of canonical analysis: canonical regression, redundancy analysis and partial and bipartial canonical analysis will also be discussed.

2.1. Canonical correlation analysis

In studying the association between two sets of variables, the researcher's interest may be focused on the degree of association between the sets. In such a case it is sufficient to know the significant canonical correlations between both sets of variables. When testing, the assumptions of multivariate normality and homogeneity of variance have to be fulfilled. In generalizing the results from a canonical correlation analysis one has to realize that canonical correlation capitalizes on sample specific error. So the weights found may be at variance with the population weights. As with most multivariate analyses it is recommendable to cross-validate the results. This can be achieved by a split half of the sample. The canonical weights can be computed on one half of the sample. By using these weights to compute on the other half the variate scores, canonical correlations and loadings, a better insight into the error may be gained. This cross-validation procedure is similar to those followed in other multivariate analyses. A third recommendation concerns the number of variables in the final solution. Usually a large number of variables show low correlations with the canonical variates. It is then appropriate to select (*e.g.* after rotation) the most important variables to represent the predictor and criterion sets for further interpretation.

2.2. Canonical regression analysis

In the foregoing, following the conventions in the literature, a distinction is made between predictor and criterion variables mainly to distinguish between the two sets of variables. In canonical regression this distinction implies a causality. The predictor set contains the explaining variables and the criterion set contains the variables to be explained, similar to common regression analysis. In canonical regression analysis as opposed to common regression we do not have one observed variable to explain but more composite variables, the criterion variates. The advantage of canonical regression analysis over common regression analysis is that more than one criterion variable can be included in the analysis. This actually means that the interrelationships between criterion variables are taken into account. As there is no single criterion variable, the concepts of error and explained variance can be troublesome, because they pertain to unobserved canonical variates.

Thus, an attempt to account for the omitted variance from the original variables is problematic. But as we want to incorporate the interrelationships between the criterion variables it may not be worthwhile to bother about the explained variance. In fact, we may want to incorporate more criterion variables because reality is complex and we hope that a composite criterion variate may better reflect reality than single criterion variables in separate regression analyses. Thus when a canonical regression analysis has been performed, the first step is to interpret the canonical loadings of the criterion variates found before or after rotation. If the criterion variates reveal interpretable underlying constructs we proceed by regressing the criterion variates on the predictor variables similar to common regression analysis. The canonical regression weights can be found from formula (4).

$$\hat{Y}_1 = \beta_1 X_1 + \dots \beta_m X_m + e = (W_{x1} \cdot R_{c1}) X_1 + \dots + (W_{xm} \cdot R_{c1}) X_m + e \quad (4)$$

In which: $\hat{Y}_1$ = the estimate of the first criterion variate;
X_1 to X_m = the predictor variables;
β_1 to β_m = the canonical regression weights;
W_{X1} to W_{xm} = the canonical correlation weight for variables X_1 to X_m on the first predictor variate;
R_{c1} = the first canonical correlation;
e = error term.

As is shown in formula (4) the canonical regression weights can be found by multiplying the canonical correlation weights from the predictor variate (W_{x1}) with the canonical correlation of the corresponding variate (R_{c1}). These beta-weights may be interpreted similar to beta-weights in common regression analysis.

2.2.1. The evaluation of canonical regression results

As stated before, the first step is to evaluate the criterion variate based on the loadings of criterion variables. When the criterion variates reveal well-interpretable underlying constructs with a sufficient amount of explained within-set variance, according to formula (1) two evaluation criteria for the canonical regression results are available:

(1) The eigenvalue R_c^2 of the canonical variate
(2) The redundancy index from Stewart and Love (1968).

The R_c^2 is the squared canonical correlation between a pair of predictor and criterion variates. The R_c^2 expresses the amount of variance explained from the criterion variate by the predictor variate. A disadvantage of this measure is that it does not incorporate an evaluation of the usefulness of the criterion variate. It has to be complemented with an evaluation of the usefulness of the criterion variate, with the amount of within (criterion) set variance explained. The R_c^2 will in most cases be the least attractive as it is a consequence of the canonical correlation procedure in which a maximum relationship is created between constructed variables, the canonical variates, without taking the loadings of observed variables into account. When one is interested in the explanatory power of a set of predictors for the observed criterion variables and not for the constructed variates, the R_c^2 may be inflated (too high). In these cases the researcher will have to inspect the loadings and explained within-set variance for the predictor and criterion variates.

The redundancy index of Stewart and Love (1968) (see formula 2) is a measure which explicitly takes the explained variance of the observed variables into account. The redundancy index provides a summary measure of the average ability of a set of predictor variables to explain variations in a set of criterion variables. Compared to R_c^2 the redundancy measure is a less inflated measure of the magnitude of the relationships. However, when there is a high within set variance in the criterion variates, with high canonical correlations between sets and the criterion variates are poorly explained by a great number of predictor variables, the redundancy index may be misleading.

2.2.2. *Cross loadings*

A cross loading expresses the relationship between an observed variable from one set with a canonical variate from the other set. The advantage of the cross loadings for instance of criterion variables with the predictor variates, is that they express the relationship of each variable separately with the predictor variate without interference of other predictor variables. The cross loadings are more conservative, less inflated than within-set loadings and form a more solid base for interpretation. The sum of the squared cross loadings in a set divided by the corresponding number of variables gives the redundancy coefficient. This procedure is an alternative for the calculation from formula (2).

Table 4
Canonical correlation *versus* redundancy analysis.

Variables	Variate 1 can. loadings		Variate 2 can. loadings	
Predictors	CC [a]	CR [b]	CC	CR
Extensive	0.53	0.47	0.35	0.39
Sober	−0.43	−0.49	0.00	−0.07
Imaginative	0.50	0.55	0.24	0.19
Varied	0.69	0.72	0.15	0.07
With family	0.14	0.11	0.53	0.59
With care	0.24	0.26	0.58	0.45
Fast	−0.26	−0.21	−0.43	−0.39
Neat	0.08	0.04	0.66	0.65
Good looking	0.26	0.22	0.57	0.64
Explained variance	15.7	14.6	19.6	17.3
Criteria				
Bread	0.48	0.40	0.12	0.24
Meats	0.72	0.74	0.06	0.07
Cheese	0.45	0.53	0.21	0.12
Eggs	0.39	0.44	0.30	0.25
Table cloth	0.13	0.17	0.72	0.72
Dishes	0.05	0.05	0.46	0.39
Explained variance	18.7	20.2	14.7	13.5
Canonical correlation	0.95	–	0.63	–
Redundancy	17	20.2	9	13.5

[a] CC: the results from the canonical correlation analysis.
[b] CR: the results from the redundancy analysis.

2.3. Redundancy analysis

In canonical correlation the finding of a maximum correlation between the two sets of variables is emphasized disregarding the explained variance in the observed variables. Rotation and further calculations as mentioned above may give complementary information. However, when we want to lay stress on the maximization of the explained variance in the criterion variables, a good alternative can be found in redundancy analysis (Wollenberg 1977).

Redundancy analysis maximizes the redundancy index of Stewart and Love (1968). In redundancy analysis it is not necessary to determine variates from both sets. If, based on theoretical insights, the variables can be subdivided in a criterion and a predictor set, the predictive qualities of the predictors for the criterion variables may be determined without incorporating the criterion part. This has important advantages for the interpretation and the computation of variate scores as elimination of variables from one set has less effect on the other set.

In table 4 a comparison is made of a redundancy analysis and the canonical correlation analysis on our example of tables 1 and 2. From table 4 we see that the canonical correlation solution succeeds in explaining resp. 17 and 9 percent of the variance in the criterion set, while the redundancy analysis explains 20 and 13.5 percent of the variance in the criterion set. As redundancy analysis is a non-symmetric form of analysis, a canonical correlation coefficient is not relevant there.

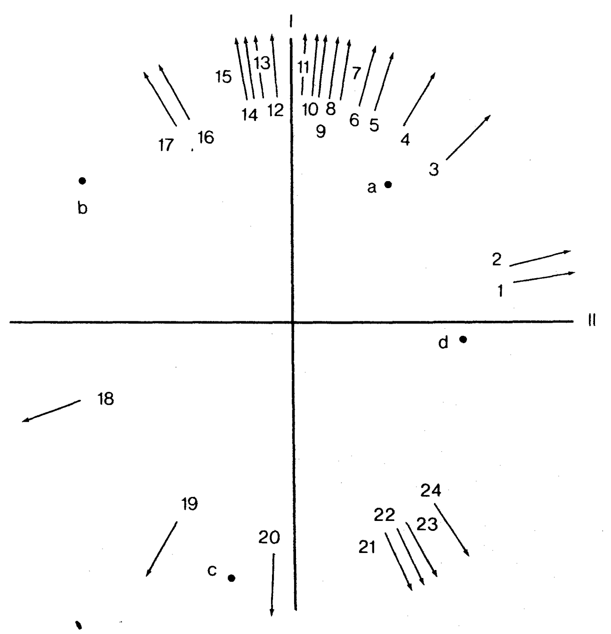
2.4. Partial and bipartial canonical correlation

In some cases the relationship between two variables is influenced by a third variable. In order to get a clear picture of the relationship between the two variables the influence of the third variable has to be eliminated. When a researcher is confronted with large data sets, the problem of eliminating the effect of intervening variables is much more complicated. Then the effect on the relationship between two other sets of variables (canonical correlation) of a set of variables should be eliminated instead of a single variable. Partial canonical correlation is the ordinary canonical correlation between two sets of variables X and Y after eliminating the effect of a third set of variables, Z. If, however, Z does not influence the variation in both X and Y, the best procedure is to partial out the influence of Z on only one of the sets X and Y (bipartial canonical correlation). It is also possible to partial out the influence of a set Z from A and another set N from Y. The concepts partial and bipartial correlation offer the opportunity to use canonical correlation analysis in a stepwise fashion. Suppose we have three predictor sets X_1, X_2 and X_3 and criterion set Y and we want to know if each of three sets accounts for substantial canonical correlation, then it is possible to enter X_1 first in the analysis partialling out X_2 and X_3 from X_1, and proceed as in ordinary stepwise regression. For a more detailed discussion see Cooley and Lohnes (1971) and Timm and Carlson (1976).

3. Brand positioning using canonical analysis

The applications of canonical analysis in consumer research may be directed in finding:

- The canonical correlations, indicating the degree of association between two sets of variables *e.g.* attitudes and behavior.
- The canonical weights indicating the relevance of variables within one set in obtaining a maximum correlation with the other set.
- The canonical variate scores, expressing the scores of respondents on



a, b, c, d: the four brands

1 = playful	7 = aggressive	13 = noisy	19 = expert
2 = exciting	8 = active	14 = cheerful	20 = intelligent
3 = sporty (sporting)	9 = masculine	15 = warm	21 = exclusive
4 = childish	10 = strong	16 = honest	22 = rich
5 = adventurous	11 = young	17 = nice	23 = progressive
6 = sturdy	12 = sociable	18 = grown-up	24 = sly

Fig. 1. Brand positioning using canonical correlation.

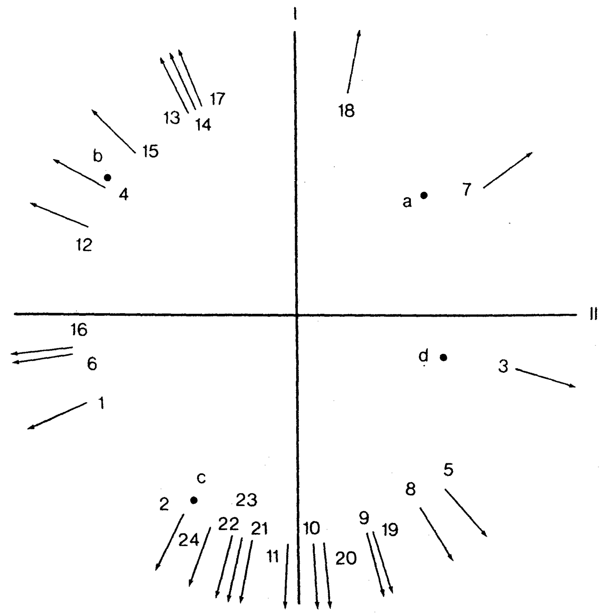
the underlying constructs, the variates. These scores are then used in cluster analysis to find market segments.

The example presented below may be seen as complementary to this latter approach. Besides the structuring of the consumer side of the market the brand side may also be structured when using canonical analysis. As a part of a larger research project sponsored by FHV/B-BDO Advertising Agency on the measurement of self images and brand images (see Verhallen and Stalpers 1980), four newly conceived brands of cigarettes were studied. The four brands were judged by 96 smokers on 24 five-point Likert items. With these data there are various ways to position the brands in a multidimensional space, identified by the Likert items. For instance, principal components analysis on the Likert scores or similarity scaling on derived brand similarities might achieve this task. However, a technical problem of obtaining a sufficiently determined structure arises when scaling only four brands. As pointed out by Huber and Holbrook (1979) principal components analysis offers a drawback because of the danger of affect-loaden dimensions (halo effects). For this reason we preferred to use canonical analysis.

Discriminant and canonical analysis tend to emphasize dimensions that are more homogeneous with respect to perception across subjects. That means that discriminant and canonical analysis tend to provide "objective" dimensions representing characteristics on which consumers agree about the positioning of brands. For a more detailed discussion of this topic the reader is referred to Huber and Holbrook (1979) and Hauser and Koppelman (1979).

3.1. Canonical analysis versus discriminant analysis

A canonical correlation analysis with the Likert scores as predictors and the K brands transformed into $K-1$ dummies as criterion variables, has been performed. By computing the scores of the brands on the first two canonical criterion variates the positions of the brands are determined as presented in fig. 1. The correlations of the Likert items with the criterion variates are the vectors of these items projected on the axes. In principle we could have used discriminant analysis to realize this end. If the criteria variables, the brands, are transformed into $K-1$ dummy variables, canonical correlation analysis is in fact the same as discriminant analysis. So discriminant analysis can be seen as a special



a, b, c, d: the four brands

1 = playful	7 = aggressive	13 = noisy	19 = expert
2 = exciting	8 = active	14 = cheerful	20 = intelligent
3 = sporty (sporting)	9 = masculine	15 = warm	21 = exclusive
4 = childish	10 = strong	16 = honest	22 = rich
5 = adventurous	11 = young	17 = nice	23 = progressive
6 = sturdy	12 = sociable	18 = grown-up	24 = sly

Fig. 2. Brand positioning using canonical dummy analysis.

case of canonical correlation analysis. Bartlett (1938) was the first to introduce multiple discriminant analysis in this way.

3.2. Optimal scaling

An advantage of canonical correlation analysis and discriminant analysis is that it offers the possibility of optimal scaling of the attribute

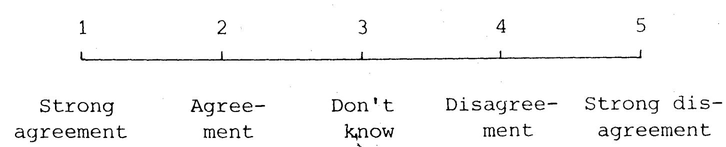


Fig. 3.

Table 5

Likert item 1	D_1	D_2	D_3	D_4
Strong agreement	1	0	0	0
Agreement	0	1	0	0
Don't know	0	0	1	0
Disagreement	0	0	0	1
Strong disagreement	0	0	0	0

categories, the position on the Likert scales, when applied to two sets of dummy variables as suggested by Green *et al.* (1978).

Starting from the methodological problem that assigning numbers to scale positions is an arbitrary decision, optimal scaling is used to maximize the relationship between the observations and the data analysis model, while respecting the measurement character of the data (Young *et al.* 1976). For example, the scale (see fig. 3) used in this study has five equidistant positions. The problem is, however, that in reality the scale positions are not equidistant. There is a possible range of values around each integer of the scale. Optimal scaling does not require the assignment of arbitrary values to the scale positions in advance. The values are determined in the analysis. This results in a better fit of these calculated scale values with the underlying attribute categories then for *a priori* assigned scale values.

The first step in obtaining these optimal scale values is to replace the five scale position of each Likert item by dummy codes, as shown in table 5.

So each Likert item is transformed into $K-1$ dummy variables. A canonical correlation analysis on the 96 dummy predictor variables, created from the original 24 Likert items, with 3 ($4-1$ brands) dummy criterion variables has been performed. The resulting first linear variate accounts for 31 percent of the variance. In substituting the canonical weights for the dummy variables we obtain the optimal scale values. Instead of the original values 1, 2, 3 and 4 we find the new scale values for example for item 1: 0.18, 0.01, -0.02 and -0.09 and for item 23: -0.31, -0.21, -0.13 and 0.00. So the equidistance of the original scale positions for item 1 is found to be inappropriate, however, for item 23 the new scale values are about equidistant. By correlating the

[2] The authors wish to thank Chris Middeldorp for his assistance in the analysis for this example.

rescaled items with the criterion variates we obtain the highest possible correlations between the predictors and the criterion variates. The results of this analysis [2] is depicted in fig. 2. A sharper portrayal of the four brands by the Likert items as compared with fig. 1 is the result: Rao's generalized distance for the four brands in the canonical analysis is 5.90 (for two dimensions) and 7.75 in the canonical dummy analysis.

References

Alpert, M.J., 1971. A canonical analysis of personality and the determinants of automobile choice. Combined proceedings. Chicago: American Marketing Association.

Alpert, M.J. and R.A. Peterson, 1972. On the interpretation of canonical analysis. Journal of Marketing Research 9, 187–192.

Bartlett, H.S., 1938. Further aspects of multiple regression. Proceedings of the Cambridge Filosofic Society 34, 34–40.

Baumgarten, S.A. and L.W. Ring, 1971. An evaluation of media readership constructs and audience profiles by use of canonical correlation analysis. Combined proceedings. Chicago: American Marketing Association.

Carmone, F.J., 1977. 'Canonical analysis in marketing research'. In: J.N. Sheth (ed.), Multivariate methods for market and survey research. Chicago: American Marketing Association.

Cooley, W.W. and P.R. Lohnes, 1971. Multivariate data analysis. New York: Wiley.

Darden, W.R. and F.D. Reynolds, 1971. Shopping orientations and product usage rates. Journal of Marketing Research 8, 505–508.

Fornell, C. and R.A. Westbrook, 1978. Identification of consumer information gathering approaches: application of a functional/structural segmentation methodology. Paper presented at the annual meeting of the European Academy for Advanced Research in Marketing, Stockholm.

Frank, R.E. and C.E. Strain, 1972. A segmentation research design using consumer panel data. Journal of Marketing Research 9, 385–390.

Green, P.E., V.R. Rao and W.S. DeSarbo, 1978. Incorporating group-level similarity judgements in conjoint analysis. Journal of Consumer Research 5, 187–193.

Hauser, J.R. and F.S. Koppelman, 1979. Using attribute ratings for product positioning: some distinctions among compositional approaches. Journal of Marketing Research XVI (November), 507–516.

Huber, J. and M.B. Holbrook, 1979. Alternative perceptual mapping techniques: relative accuracy and usefulness. Journal of Marketing Research XVI (November), 495–506.

Johnston, J., 1972. Econometric methods. New York: McGraw-Hill.

Johansson, J.K. and C. Lewis, 1974. A clarification of the redundancy index. Working Paper, College of Commerce and Business Administration. University of Illinois, Urbana, IL.

Nicewander, W.A. and D.A. Wood, 1974. Comments on a general canonical correlation index. Psychological Bulletin 81, 92–94.

Schaninger, C.M., V. Parker Lessig and D.B. Panton, 1980. The complementary use of multivariate procedures to investigate non-linear and interactive relationships between personality and product usage. Journal of Marketing Research 17, 119–124.

Sparks, D.L. and W.T. Tucker, 1971. A multivariate analysis of personality and product use. Journal of Marketing 8, 67–70.

Stewart, D. and W. Love, 1968. A general canonical correlation index. Psychological Bulletin 70, 160–163.

Timm, N.H. and J.E. Carlson, 1976. Part and partial correlation analysis. Psychometrika 41, 159–175.

Verhallen, Th.M.M. and J. Stalpers, 1980. The self-concept and brand choice. Paper presented at the Fifth Colloquium on Economic Psychology. Leuven, Brussels.

Wendt, J.C., 1979. Canonical correlation as an explanatory technique of attitude scale construction. Public Opinion Quarterly.

Wollenberg, A.L. van den, 1976. Redundancy analysis. An alternative for canonical correlation analysis. Psychometrika 41, 207–219.

Young, F.W., J. de Leeuw and Y. Takana, 1976. Regression with qualitative and quantitative variables: an alternating least squares method with optimal scaling features. Psychometrika 41, 505–529.